FILENAME.APP=4106MS00.DOC

IMAGE PROCESSING
Software: MS-Word97 for Windows standard edition.
25 Pages

Dr. Harpreet Singh and Y. Hamzeh
Department of Electrical and Computer Engineering
Wayne State University
Detroit MI 48202

Dr. S. Bhama
Client Server System Laboratory
JTICS, Chrysler Corporation
Detroit MI 48227

Dr. S. Talahmeh
INVANTAGE, Inc.
Taylor MI 48180

Dr. L. Anneberg
Department of Electrical Engineering
Lawrence Technological University
Southfield MI 48015

Dr. G. Gerhart and Dr. T. Meitzler
U.S. Army Tank Automotive and Armaments Command, Survivability Technology Center,
Research and Engineering Center (TARDEC), Warren MI 48092

Dr. D. Kaur
Department of Electrical Engineering and Computer Sciences
University of Toledo
Toledo OH 43606

Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **01 MAR 1999**	2. REPORT TYPE **Technical Report**	3. DATES COVERED **01-03-1999 to 01-03-1999**
4. TITLE AND SUBTITLE **Image Processing**		5a. CONTRACT NUMBER
		5b. GRANT NUMBER
		5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S) **Thomas Meitzler; Grant Gerhart; Harpreet Singh; S Bhama**		5d. PROJECT NUMBER
		5e. TASK NUMBER
		5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **U.S. Army TARDEC ,6501 E.11 Mile Rd,Warren,MI,48397-5000**		8. PERFORMING ORGANIZATION REPORT NUMBER **#13905**
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) **U.S. Army TARDEC, 6501 E.11 Mile Rd, Warren, MI, 48397-5000**		10. SPONSOR/MONITOR'S ACRONYM(S)
		11. SPONSOR/MONITOR'S REPORT NUMBER(S) **#13905**

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
Image processing consists of a wide variety of techniques and mathematical tools to process an input image. An image is processed as soon as we start extracting data from it. The data of interest in object recognition systems are those related to the object under investigation. An image usually goes through some enhancement steps, in order to improve the extractability of interesting data and subside other data. Extensive research has been carried out in the area of image processing over the last 30 years.Image processing has wide area of applications. Some of the important areas of application are business, medicine, military and automation. Image processing has been defined as a wide variety of techniques that includes coding, filtering,enhancement, restoration registration, and analysis. In many applications, such as the recognition of three-dimensional objects, image processing and pattern recognition are not separate disciplines. Pattern recognition has been defined as a process of extracting features and classifying objects. In every 3D-object recognition system there are units for image processing and there are others for pattern recognition. There are three different approaches to image processing: a. Analog processing: This approach is very fast since the time involved in AD and a DA conversion is saved. But this approach is not flexible since the manipulation of images is very hard. b. Digital processing: This approach is slower than the analog approach but is very flexible, since manipulation is done very easily. The processing time of this approach is tremendously improved by the advent of parallel processing techniques. c. Heuristic processing: This is a smart way of processing but it is conditional since it involves probabilities. The dominant approach is the digital processing approach because it is very flexible.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**59**	

Image Processing

Image processing consists of a wide variety of techniques and mathematical tools to process an input image. An image is processed as soon as we start extracting data from it. The data of interest in object recognition systems are those related to the object under investigation. An image usually goes through some enhancement steps, in order to improve the extractability of interesting data and subside other data.

Extensive research has been carried out in the area of image processing over the last 30 years. Image processing has wide area of applications. Some of the important areas of application are business, medicine, military and automation.

Image processing has been defined as a wide variety of techniques that includes coding, filtering, enhancement, restoration registration, and analysis. In many applications, such as the recognition of three-dimensional objects, image processing and pattern recognition are not separate disciplines. Pattern recognition has been defined as a process of extracting features and classifying objects. In every 3D-object recognition system there are units for image processing and there are others for pattern recognition.

There are three different approaches to image processing:

  a. Analog processing: This approach is very fast since the time involved in AD and a DA conversion is saved. But this approach is not flexible since the manipulation of images is very hard.

  b. Digital processing: This approach is slower than the analog approach but is very flexible, since manipulation is done very easily. The processing time of this approach is tremendously improved by the advent of parallel processing techniques.

  c. Heuristic processing: This is a smart way of processing but it is conditional since it involves probabilities. The dominant approach is the digital processing approach because it is very flexible.

Digital Image Processing

Digital image processing is defined in [1] as "the processing of two dimensional images by a digital computer". This implies that the image data must be converted to the binary form, the only form understood by computers. Digital image is represented by an array of regularly spaced and very small-quantized samples of the image. Two processing that are related to any digital system are the sampling and quantization.

When a picture is digitized, it is represented by regularly spaced samples of this picture. These quantized samples are called pixels. The array of pixels that are processed in practice can be quite large. To represent an ordinary black and white TV image digitally, an array of (512 x 512) pixels is required. Each pixel is represented by an 8-bit number to allow 256 gray levels. Hence a single TV picture needs about 2 million bits.

Digital image processing encompasses a wide variety of techniques and mathematical tools. They have all been developed for use in one or the other of the two basic activities that constitute digital image processing; image preprocessing and image analysis.
An approach called state-space approach has been recently used in modeling image processors. These image processors are made of linear iterative circuits. The state-space model is used efficiently in image processing and image analysis. If the model of an image processor is known, the realization of a controllable and observable image processor is then very simple.

Image Preprocessing

Image preprocessing is an early stage activity in image processing that is used to prepare an input image for analysis to increase its usefulness. Image preprocessing includes image enhancement, restoration and registration.
Image enhancement accepts a digital image as input and produces and enhanced image as an output; in this context, enhanced means better in some respects. This includes improving the contrast, removing geometric distortion, smoothing the edges, or altering the image to facilitate the interpretation of its information content.
In image restoration, the degradation is removed from the image to produce a picture that resembles the original undegraded picture.
In image registration, the effects of sensor movements are removed from the image or to combine different pictures received by different sensors of the same field.

Image Analysis

Image analysis accepts a digital image as input and produces data or a report of some type. The produced data may be the features that represent the object or objects in the input image. To produce such features, different processes must be performed that include segmentation, boundary extraction, silhouette extraction and features extraction. The produced features may be quantitative measures, such as moment invariants, Fourier descriptors or even symbols, such as regular geometrical primitives.

Sampling and Quantization

Quantization is the process of representing a very large number (possibly infinite) of object with a smaller, finite number of objects. The "representing" set of objects may be taken from the original set (e.g. the common number rounding process) or may be completely different (e.g. the alphabetical grading system commonly used to represent test results).
In image processing systems, quantization is preceded by another step called sampling. The gray level of each pixel in an image is measured, and a voltage signal that is proportional to the light intensity at each pixel is generated. It is clear that the voltage signal can have any value from the voltages that are generated by the sensing device.

Sampling is the process of dividing this closed interval of continuous voltage signal into a number of subintervals that are usually of equal length. In an 8-bit sampling/quantization process, for example, the interval of voltage signals is divided into 256 subintervals of equal length.

In the quantization process, each of the generated intervals from sampling is represented by a codeword. In an 8-bit quantization process, each codeword consists of 8 bits binary number. An 8 bit Analog-to-Digital Converter (ADC) can simply accomplish the tasks of sampling and quantization. The image data is now ready for further processes using digital computers.

For systems that involve dynamic processing of image signals (e.g. TV signals or video streams from CCD cameras), the term sampling refers to a completely different process. In this context, sampling means taking measurements of the continuous image signal at different instants of time. Each measurement can be thought of as a single stationary image. A common problem associated with image digitization is aliasing.

The sampling theorem states that for a signal to be completely reconstructable, it must satisfy the following equation:

$$w_s = 2 \cdot w \tag{1.1}$$

Where w_s is the sampling frequency and w is the frequency of the sampled signal. Sampling, in this context, means taking measures of the analog signal at different instants, separated by a fixed time interval Δt. This theorem is applicable on the sampling of stationary images as well, where sampling is carried through space instead of time. If the signal is bandlimited, sampling frequency is determined according to the frequency of its highest frequency component. Image signals, however are subjected to truncating, mainly because of the limitations in sensors and display devices. Sensors are capable of recognizing a limited range of gray level. Real objects usually have wider ranges of gray levels, which means that both the gray levels higher and lower than the range of the sensor are truncated. Truncating is what causes the aliasing problem. To explain how this happens, consider the simple sinusoidal function given by:

$f(x) = cos(x)$. Figure (1) shows a plot of this function and figure (2) shows a plot of its Fourier transform. Figure (3) shows a truncated version of that function, and figure (4) shows the equivalent Fourier transform. This function has infinite duration in the frequency domain. The Nyquist frequency is given by: $w_n = w_s / 2$, where w_s is the sampling frequency If we try to sample this signal with a sampling frequency of w_s, then all frequencies higher than the Nyquist frequency will have aliases within the range of the sampling frequency. In other words, aliasing causes high frequency components of a signal to be seen as low frequencies.

This is known as folding. A practical method to get red of aliasing is to prefilter the analog signal before sampling. Figure (4) shows that lower frequencies of the signal contains most of signal's power. A filter is designed so that filtered signals do not have frequencies above the Nyquist frequency. A standard analog filter transfer function may be given as:

$$G_f(s) = \frac{w^2}{s^2 + 2\zeta w s + w^2} \tag{1.2}$$

Where ζ is the damping factor of the filter and w is its natural frequency. By cascading second and first order filters, one can get higher order systems that have higher performances. Three of the most commonly used filters are the Butterworth filter, ITAE filter and Bessel filter. Bessel filters are commonly used for high performance applications, mainly because of the following two factors:

1. The damping factors that may be obtained by a Bessel filter are generally higher than those obtained by other filters. A higher damping factor means a better cancellation of frequencies outside the desired bandwidth.

2. The Bessel filter has a linear phase curve, which means that the shape of the filtered signal is not much distorted.

To demonstrate how we can use Bessel to get rid of high-frequency noise and anti-aliasing, consider the square signal in figure (5). This has a frequency of 25 Hz. Another signal with a frequency of 450 Hz is superimposed on the square signal. If we try to sample the square signal with noise (figure 5-a), we will get a very distorted signal (figure 5-b). Next, we prefilter this signal using a second-order Bessel filter with a bandwidth of 125 Hz and a damping factor of 0.93. The resultant signal is shown in figure (5-c). Figure (5-d) shows the new signal after Sampling. It is clear that this signal is very close to the original one square signal without noise.

Image restoration

Image restoration refers to a group of techniques that are oriented toward modeling the degradation and applying the inverse process in order to recover the original image.

Each component in the imaging system contributes to the degrading of the image. Image restoration techniques try to model the degradation effect of each component and then perform operations to undo the model, to restore the original image [3].

There are two different modeling approaches for degradation: the priori approach and the posteriori approach. These two approaches differ in the manner in which information gathered to describe the characteristics of the image degrading.

The priori approach is to try to model each source of noise in the imagery system by measuring the system's responses to arbitrarily noises. In many cases, deterministic models can not be extracted and stochastic models are used instead.

The posteriori approach is adopted when a great deal of information is known about the original image. We can develop a mathematical model for the original image and try to fit the model to the observed image. Figure (6) shows a simplified model for the image degradation and restoration processes. The original image signal $f(x, y)$ is subjected to the linear degrading function $h(x, y)$. An arbitrarily noise signal $\eta(x, y)$ is then added, to create the degraded image signal $g(x, y)$. Reconstruction approaches try to estimate the original image signal $f(x, y)$ given the degradation signal $g(x, y)$ and some statistical knowledge of the noise signal $\eta(x, y)$.

We can broadly classify reconstruction techniques into two classes, the filtering reconstruction techniques and the algebraic techniques [5].

Filtering reconstruction techniques

These techniques are rather classical and they make use of the fact that noise signals usually have higher frequencies than image signals. This means that image signals die out faster than noise signals in high frequencies. By selecting the proper filter, one can get a good estimate of the original image signal, by reducing the effect of noise. Examples of the reconstruction filters are the deconvolution filter and the Wiener filter.

Deconvolution filter

This filter is based on the concept of inverse filtering, where the transfer function of the degraded system is inverted to produce a restored image. Figure (7) shows typical spectra of deconvolution filtering image restoration system. If no noise is added to the system, the image is perfectly reconstructed. The presence of noise, however, will add reconstruction error, whose value can become quite large at spatial frequencies for which $f(x, y)$ is small. This reflects directly on regions of the image with high details because they have higher spatial frequencies than other regions.

Wiener filter

This filter uses the mean-squared error (MSE) criterion to minimize the error signal between the original and degraded image signals. The mean-squared restoration error is given by:

$$\varepsilon = E\{[f(x,y) - g(x,y)]^2\}$$ (2.1)

Where $f(x, y)$ and $g(x, y)$ are as before. The Wiener filter can be given by:

$$W(m,n) = \frac{H*(m,n)}{|H(m,n)|^2 + P_n(m,n) / P_f(m,n)}$$ (2.2)

Where $P_n(m,n)$ and $P_f(m,n)$ are the power spectra of the signal and noise, respectively.
Wiener filter acts as a band-pass filter. At low spatial frequencies, it acts as an inverse filter, whereas at higher frequencies, is acts as a smooth rolloff low-pass filter.
This filter is not very suitable for use in the cases where images are investigated by the human eye. The MSE technique treats all errors equally, regardless of their spatial location in the image. The human eye, on the other hand, has high degree of tolerance to errors in darker areas on the image than elsewhere. Another limitation in this filter is that it can not handle dynamically changing image and noise signals.
Another type of filter that is closely related to the Wiener filter is the power spectrum equalization filter, which can be given by:

$$PSE(m,n) = \left[\frac{1}{|H(m,n)|^2 + P_n(m,n) / P_f(m,n)}\right]^{1/2}$$ (2.3)

The power spectrum obtained by this filer for the reconstructed filter is identical to that of the original signal.

Linear algebraic restoration techniques

These techniques utilize matrix algebra and discrete mathematics for solving the problem of image restoration. To extract the discrete model for the restoration system, we make the following assumptions:

- The digitized original image $f(x, y)$ and the restored image $f_r(x, y)$ are stored in the M^2-by-1 column vectors f and f_r, respectively.
- The digitized degrading function $h(x, y)$ is stored in a square matrix H (M^2-by-M^2).

- The degraded image $g(x, y)$ and noise $\eta(x, y)$ is stored in the M^2-by-1 column vectors **g** and **n**, respectively.

We can express the observed (degraded) image vector in the compact form:

$$\mathbf{g} = \mathbf{Hf} + \mathbf{n} \tag{2.4}$$

We use this model to derive some of the algebraic restoration techniques. Namely the unconstrained reconstruction technique, the constrained reconstruction technique and the pseudoinverse filtering technique.

Unconstrained reconstruction

If we know very little about the noise **n**, then we try to find an estimate image $\mathbf{f_r}$, such that $\mathbf{Hf_r}$ approximates **g** in a least-squares manner. This can be accomplished by minimizing the norm of the noise **n**. Squaring the norm of both sides of equation (3.4) after substituting for **f** by the estimate vector $\mathbf{f_r}$ and moving $\mathbf{Hf_r}$ to the other side of the equation, we get:

$$\|\mathbf{n}\|^2 = \|\mathbf{g} - \mathbf{Hf_r}\|^2 = (\mathbf{g} - \mathbf{Hf_r})(\mathbf{g} - \mathbf{Hf_r})' \tag{2.5}$$

Where $\|\mathbf{a}\|^2$ is the square of the norm of vector **a** and is given by $\|\mathbf{a}\|^2 = \mathbf{aa}'$, where a' is the transpose of the vector a. Consider the error function E, where:

$$E(\mathbf{f_r}) = \|\mathbf{g} - \mathbf{Hf_r}\|^2 = (\mathbf{g} - \mathbf{Hf_r})(\mathbf{g} - \mathbf{Hf_r})' \tag{2.6}$$

Then our goal is to minimize E with respect to $\mathbf{f_r}$. This can be accomplished by taking the derivative of E with respect to $\mathbf{f_r}$, and equating the result to zero, that is:

$$\frac{\partial E(\mathbf{f_r})}{\mathbf{f_r}} = -2\mathbf{H}'(\mathbf{g} - \mathbf{Hf_r}) = 0 \tag{2.7}$$

Assuming that $\mathbf{H}^{-1}$ exists, the solution for this equation is given as:

$$\mathbf{f_r} = \mathbf{H}^{-1}\mathbf{g} \tag{2.8}$$

Constrained reconstruction

To account for the noise term in equation (2.4), we introduce the square matrix **Q** (M^2-by-M^2) to represent some linear operator on f. By selecting different **Q**'s, we are able to set the goal of restoration as desired. Equation (2.6) is now modified to:

$$E(\mathbf{f_r}) = \|\mathbf{Qf_r}\|^2 + \lambda(\|\mathbf{g} - \mathbf{Hf_r}\|^2 - \|\mathbf{n}\|^2) \tag{2.9}$$

Where λ is a constant called the LaGrange multiplier. Again we try to minimize the error function E, by taking its derivative with respect to $\mathbf{f_r}$, and equating the result to zero, that is

$$\frac{\partial E(\mathbf{f_r})}{\mathbf{f_r}} = 2\mathbf{Q}'\mathbf{Qf_r} - 2\lambda\mathbf{H}'(g - \mathbf{Hf_r}) = 0 \tag{2.10}$$

Solving for $\mathbf{f_r}$, we get:

$$\mathbf{f_r} = (\mathbf{H}'\mathbf{H} + \alpha\mathbf{Q}'\mathbf{Q})^{-1}\mathbf{H}'\mathbf{g} \tag{2.11}$$

Where $\alpha = 1/\lambda$ is a constant that we adjust to satisfy the constraint.

Pseudoinverse filtering

The pseudoinverse filtering is a special case of the constrained reconstruction techniques. In this technique, the constraint matrix **Q** is the M^2-by-M^2 identity matrix. Equation (2.11) then becomes:

$$f_r = (H'H + \alpha I)^{-1}H'g \qquad\qquad (2.12)$$

If $\alpha = 0$, this equation reduces to equation (2.8), which represents the unconstrained restoration technique.

Image Enhancement

Image enhancement refers to a group of processes that aim toward making the processed image much better than the unprocessed image. In this sense, image enhancement is closely related to image restoration. Enhancement techniques are basically heuristic procedures, which are designed to manipulate an image, by taking advantage of the human vision system [2].

Image enhancement is used to remove the noise that may be added to the image from several sources including electrical sensor noise and transmission channel errors. There are many enhancement techniques to improve the input image. Some of these techniques are gray scale modification, point operators, contrast enhancement and filtering.

Gray Scale Modification

The gray-scale modification technique is very effective to modify contrast or dynamic range of an image. In this technique, the gray scale of an input image is changed to a new gray scale according to a specific transformation. The specific transformation is the relation between the intensity of the input image to the intensity of the output image. By properly choosing this transformation, one can modify the dynamic range or the contrast of an input image. For each input pixel the corresponding output pixel is obtained from the plot or the look-up table.

The transformation desired depends on the application. In many applications a good transformation is obtained by computing the histogram of the input image and studying its characteristics. The histogram of an image represents the number of pixels that has a specific gray value as a function of all gray values.

The gray-scale transformation is very simple and very efficient but it has the following two limitations:

      1. It is global: The same operation is applied to all the pixels that have the same gray-level in the image.

      2. The transformation that is good for one image may not be good for others.

Point Operators

Point processing is a group of simple techniques used for image enhancement. Point operator is a direct mapping from an input image to an output image, where the pixel value in the enhanced output image depends only on the value of the corresponding pixel in the input image [5]. This mapping can mathematically be given by:

$$Po = N(Pi) \qquad\qquad (3.1)$$

Where *Po* is the output enhanced pixel, *Pi* is the corresponding input pixel and N is the point operator. In case of discrete images, a point operator may be implemented by a *LUT* (Look-Up-Table), where for each pixel in the output image, there is a corresponding pixel in the input image. It should be noted that point operators affects the gray level of pixels in an image but has no effect on the spatial relationships between pixels.

Point operators may be used in applications such as photometric calibration, contrast enhancement, display calibration and contour lines.

Types of point operations

Point operators are generally classified as either linear or nonlinear operators.

Linear point operators: The output gray level is a linear function of the input gray level. The Mapping function is given by:

$$Go = N(Gi) = aGi + b \tag{3.2}$$

Figure (8) shows a plot of a linear point operator.

Where *Go, Gi* are the gray levels of the output pixel and input pixel, respectively. By changing the values of the coefficients *a* and *b*, different effects may be accomplished.

For example, if $a = 1$ and $b = 0$, the gray levels of the image pixels do not change. A negative value of a has the effect of complementing the image. That is, dark areas of the image become light and light areas become dark.

Nonlinear Monotonic Point Operations: Nonlinear point operations are mainly used to modify the midrange gray level.

One type of nonlinear point operators increases the gray level of midrange pixels, while leaving dark and light pixels little changed. This type of operators my be given by:

$$Go = Gi + AGi(G_m - Gi) \tag{3.3}$$

Where G_m is the maximum gray level and A is a parameter that determines the amount of increase (A > 0) or decrease (A < 0) in the midlevel gray range.

The *logarithmic* operator is a point operator that maps each pixel value to its logarithm. The logarithmic mapping function is given by:

$$Go = clog\,(Gi+1) \tag{3.4}$$

In an 8-bit format, the value of c is selected such that the maximum output value is 255.

Contrast enhancement

Contrast is defined as the range of brightness values present in image. Contrast enhancement is a very important issue when the final output image has to be evaluated by a human observer. The human eye can distinguish up to 64 of gray levels, compared to 256 different gray levels that can be achieved with an 8-bit imagery system.

The range of brightness values collected by a sensor may not match the capabilities of the output display medium. Contrast enhancement involves changing the range of brightness in an image in order to increase contrast.

Selecting the suitable method for contrast enhancement depends on two main factors: the state of the original image and the requirements from the final image. In many cases, more than one step of contrast enhancement is required to obtain the desired output. Two common contrast enhancement methods are contrast stretching technique and window-and-level technique.

Contrast Stretching

This process expands the range of brightness of an original digital photo into a new distribution. As a result, the total range of sensitivity of the display device can be utilized. Contrast stretching makes minute variations within the image data more obvious.

Contrast stretching may be accomplished by both linear and nonlinear methods. Examples of linear contrast enhancement are the Minimum-maximum linear contrast stretch, Percentage linear contrast stretch and piece-wise linear contrast stretch. An example of nonlinear contrast enhancement is the Histogram equalization.

Window and Level

If the brightness range of the element of interest in an image is limited, then a method known as Window and Level can be used to enhance its contrast on the expense of other elements. All signal levels below the desired range will be mapped onto the same low output level and all those with higher levels will be mapped onto the same high output level. Only pixels within the desired range will be represented with an acceptable range of contrast.

Two parameters define the desired range of contrast: the middle point Level and the width of the range Window. Figure (9) shows how window and level technique affects the histogram of the image.

If the original image has a poor contrast level, then the window and level alone will not produce satisfactory results. A better technique would be to use one of the contrast stretching methods to increase the contrast range first, then apply the window and level method on the desired element, taking into consideration its new range of contrast.

Low-Pass Filtering

Low-pass filtering reduces the high-frequency components while retains the low- frequency components. The low frequency components of an image constitute most of the energy in an image. In other words, low-pass filtering removes a large amount of noise at the expense of removing a small amount of the actual signal.

Low pass filtering can be used efficiently to remove additive random and multiplicative noise, but at the same time it blurs the processed image. Blurring is the primary limitation of low-pass filtering. Figure (10) shows a photo of the famous Taj-Hahal, one of the Seven Wonders of the World. Figure (11) shows the effect of low-pass filtering on the same image.

High-Pass Filtering

High-pass filtering reduces the low frequency components while preserving the high frequency components of an input image. Since the high frequency components generally correspond to edges or fine details of an image, high-pass filtering increases the contrast and thus sharpens the edges of the image.

High-pass filtering can be used to remove image blur but at the same time it tends to increase background noise power. Figure (12) shows the effect of high-pass filtering on figure (10).

Median Filtering

Median filtering is a non-linear digital technique used for image enhancement. The basic idea behind median filtering is the removing of noise by finding the median of neighboring pixels and assigning it to the center point. It is used to reduce impulsive noise and to enhance edges while reducing random noise. Figure (13) shows the effect of filtering the photo image with a median filter with a window of three pixels.

In general, there are two standard local neighborhoods, which are considered, namely, square, cross-windows. These two types are presented below.

Square Median Filter

The square median filter for each pixel is obtained by sliding a square window of (2m+1) by (2m+1) over the image and replacing the center point of the window by the median value of the pixels in the square window.

Cross-Window Median Filtering

The cross-window median for each pixel is obtained by sliding a square-cross window over the image and replacing the center Pixel of the window by the median value of the pixels in the cross-window.

The cross-window median filtering is computationally more effective than the square window filtering.

Pseudo Median Filtering

The median filter discussed above is easily implemented in sequential machines. The pseudo median filtering has been developed to approximate the median filtering. The pseudo median filtering is easily implemented on single-instruction stream multiple-data stream (SIMD) parallel processors. Hence, the computation time of the pseudo filtering is much less than the time of the median filtering.

Object Recognition

Object recognition includes the process of determining the object's identity or location in space. The problem of object or target recognition starts with the sensing of data with the help of sensors like video camera, thermal sensor, etc., and then interpret these data in order to recognize an object or objects. We can divide object recognition problem into two sub-problems; the modeling problem and the recognition problem.

Object Modeling

Modeling is the process of representing a real systems in an abstract manner, in order to study its different features. It is widely used in all fields of Engineering. In control engineering, for example, a mathematical model of a physical system is extracted, to facilitate the study of its performance under different circumstances.

In object recognition, a model is created for the object under investigation. This model is then compared to different models that are stored in a database. If the this model matched one of the available models, say, the model of object A, then the investigated model is classified as object A. In an airplanes recognition system, for example, the database contains models of different types of airplanes.

There are two types of models in object modeling systems:

1. Geometric models: These models are represented by surface or volume primitives with attributes and relationships. The primitives and their relationships can be described geometrically. These models are used for regular man-made objects because they can be described geometrically.

2. Symbolic models: In these models, a three-dimensional object is represented by a graph whose nodes are the 3D primitives and whose arcs are the attributes and relations. These models are good for representing objects that are irregular in natural.

The primitives of the geometric models depend on the application. In general, there are three classes of primitives. The three classes are:

1. Surface or boundary primitives.
2. Sweep primitives.
3. Volumetric primitives.

The boundary primitives are bounded faces of the object. The sweep primitives are general cylinders or cones. The volumetric primitives can be categorized into three types as follows:

1. Spatial primitives: They are 3D arrays of cells which may be marked as filled with matter or not.

2. Solid geometric primitives: They are regular solid primitives, which may be combined by modified versions of the Boolean set operators-union, difference, and intersection.

3. Cell primitives: These primitives have complex shapes and they can be connected together by just one combining operator (glue).

Pattern recognition

Pattern is a process to extract features and recognize objects and patterns from a given image. Usually there are two approaches to solve the problem of pattern recognition. The two approaches are the RF processing approach and the image processing approach.

In the RF approach, processing electromagnetic waves are transmitted and then reflected from the target. The reflected waves are analyzed to recognize the target. The RF processing is usually used when the target is very far or out of range of sight.

In image processing approach a picture or an image of the field is taken. The picture is then processed and analyzed to recognize the objects in the picture.

The recognition task is usually done in two phases:

　　　1. Training phase: In this phase a pictorial database or a knowledge base is created off-line, which contains data about all objects in consideration. If the recognition system is designed to recognize airplanes, for example, the database may include data about all airplanes in consideration.

　　　2. Recognition phase: In this phase, the incoming picture is first processed to enhance it. The picture is analyzed to extract the features that characterize it. Finally the features are matched with the features stored in the database to recognize the object or objects.

Wavelets

Wavelets are a relatively new area of signal processing and offer a potentially very useful approach to image analysis. Because of the fact that one has local control over resolution, many cue features of different size can be extracted from an image. Wavelets also show promise as an aid in understanding how early vision can be emulated and further understood.

Wavelet Transforms

It is well known that in signal processing, the more compact your method of representing information, the better. In the year 1822, Jean Joseph Fourier devised a very efficient way to represent the information content of a signal. His idea was to represent a signal as the sum of its frequencies. Transmitted power spectra, carrier frequencies, brain activity, NMR signals - all these global descriptions provide a lot of information in a compact manner. However, most of the power of this kind of representation vanishes when one tries to represent information that changes its nature during the course of signal recording. A good example of this kind of a signal is a musical score. A global analysis of a recording of a musical selection with a Fourier Transform (FT) will indicate what specific notes were played within the piece of music but there is no way to recover the timing of the notes. The musical score on the other hand indicates what note was played and at what time.

Wavelets represent a signal in a way such that local frequency information is available for each position within the signal. Wavelets are able to analyze a signal based on the position-varying spectra. The multi-resolution pyramidal decomposition that results is also well matched to fractals and shows great potential for the removal of background noise such as static in recordings, and for pattern recognition and texture segmentation. In the following section we discuss the conceptual understanding of how wavelets can be used in signal analysis.

A fairly old method of computing a local spectrum is to apply the FT to one specific piece of the signal at a time. This is the idea behind what is called the windowed Fourier Transform (WFT). Basically, the implementation involves using a moving rectangular piece of the whole image to isolate a portion of the signal of interest, which is then Fourier transformed. As the window slides along to different positions, the WFT gives the spectra at these positions. This kind of analysis has a fundamental problem, however, whose mathematics is, in fact, similar to the Heisenberg Uncertainty Principle in Quantum Mechanics (QM). Multiplying the signal by window function results in convoluting or mixing the signal spectrum with the spectrum of the window. Add this to the fact that as the window gets smaller, its spectrum gets wider, and we have the basic dilemma of localized spectra: the better we determine the position of the signal the poorer we localize the spectrum. This is analogous to the case in QM where increased precision in a description of, say the momentum of an electron reduces the precision available in the position of that electron. Very accurate determinations can be made or computed, but both are not available simultaneously to an unlimited degree of precision. Correspondingly, there is a fundamental physical limit to the degree of accuracy with which the frequency content can be known at a particular position and simultaneously know the position where the signal is being analyzed.

In 1946, Dennis Gabor introduced a version of the WFT that reduced this uncertainty somewhat. The Gabor transform uses a Gaussian profile for the window since the Gaussian is the function that minimizes this uncertainty. However, the underlying idea of localizing a spectrum of a signal by windowing the signal needs to be reconsidered. Obviously, care must be taken in the selection of the signal. Careful attention to the placement of the window however is not an easy task for realistic time-varying signals. We are in fact trying to do two different things at once.

Frequency is a measure of cycles per unit time or signal length. So that high frequency oscillations take much less signal length or time than do low frequency oscillations. High frequencies can be well localized in the overall signal with a short window, but low frequency localization requires a long window.

The wavelet transform takes an approach that permits the window size to scale to the particular frequency components being analyzed. Wavelets are, generally speaking, functions that meet certain requirements. The name wavelet originates from the requirement of integrating to zero by oscillating about the x-axis and being well localized. In fact, there are many kinds of wavelets. There are smooth wavelets, wavelets with simple mathematical expressions, wavelets that are associated with filters, etc. The simplest wavelet is the Haar Wavelet. Just as the sine and cosine functions are used in Fourier analysis as basis functions, wavelets are used as the basis functions for representing other functions. After the "mother" wavelet, W (x), is chosen, it's possible to make a basis out of translations and dilation of the mother wavelet. Special values for a and b such that $a=2^{(-j)}$ and $b= k^j$, where k and j are integers. This particular choice of a and b gives a so-called "sparse basis." The sparse basis choice of parameters also makes the link to multiresolution signal analysis possible.

There are many important characteristics of wavelets that make them more flexible than Fourier analysis. Fourier basis functions are localized in frequency but not in time. Small frequency changes in a FT cause changes everywhere in the time domain. Wavelets can be localized in both frequency position and in scale by the use of dilation, and in time by the use of translations. This ability to perform localization is useful in many applications. Another advantage of wavelets is the

high degree of coding efficiency, or in other words, data compression available. Many classes of functions and data can be represented by wavelets in a very compact way.

The efficient coding of information makes wavelets a very good tool for use in data compression. One of the more popular examples of this sparse coding capability is the well-known fact that FBI has standardized the use of wavelets in digital fingerprint compression. The FBI gets compression ratios on the order of 20: 1. The Image difference between the original and the wavelet stored and decompressed Image is negligible. Many more applications of wavelets currently exist. One involves the denoising of old analog recordings. For example, Coifman and his group at Yale have been using wavelets to take the static and hiss out of old recordings. They have cleaned up recordings of Brahms playing his First Hungarian Dance for the piano. Noisiness extends to the realm of data sets. Noisy data sets can be cleaned up by using wavelets. Typically the speed of wavelets is also much faster than the fastest Fast Fourier Transform (FFT). The data is basically encoded into the coefficients of the wavelets. The computational complexity of the FFT is of the order of (n log n); where n is the number of coefficients; whereas for most wavelets, the order of complexity is of the order n. figure (14) shows how wavelets are used to process data.

Many data operations, such as multiresolution signal processing, can be done by processing the corresponding wavelet coefficients.

How wavelets work is best illustrated by way of an example. The simplest of all wavelets is the Haar wavelet, shown in figure (15). The Haar wavelet, W (x) is a step function that assumes the values 1 and -1, on the scale [0,1/2] and [1/2,1]. The Haar wavelet is more than 80 years old and has been used for various applications. It can be shown that any continuous function can be approximated by Haar functions. Dilation and translations of the function $\Psi(x)$,

$$\Psi_{jk}(X) = CONST.\Psi (2^j.x- k) \tag{5.1}$$

define an orthogonal basis in the space of all square integrable functions, $L^2(R)$. This means that any element of $L^2(R)$ can be presented as a linear combination of these basis functions. The orthogonality of the wavelet pairs is checked by the following,

$$\int \Psi_{jk} \Psi_{j'k'} = \delta_{jj}\delta_{kk'} \tag{5.2}$$

Whenever $j = j'$ and $k = k'$ is not satisfied simultaneously. The constant that makes the orthogonal basis orthonormal is $2^{j/2}$.

Wavelet transforms and their use with clutter metrics

Wavelets and wavelet transforms are essentially an elegant tool from the field of mathematics that can be applied to the area of signal processing. Wavelets are used for removing noise or unwanted artifacts from images as well as acoustic data.

Clutter is a term that refers to the psychophysical task of perceiving objects in a scene while in the presence of similar objects. There are many definitions of clutter in use at the moment, however, none take into account the multiresolution capability of wavelets.

Image Coding

Image coding refers to groups of techniques that are used to represent three-dimensional (3D) object in forms manageable by computers, for the purpose of object classification and recognition. Object moments, Fourier descriptors and expert systems are examples of these techniques.

The Moments Technique

An image of a three-dimensional (3D) object may be described and represented by means of a set of its moments. The set of moments of an object can be normalized to be independent of the object primary characteristics, namely translation, rotation and scale. Hence, moments can be used to recognize 3D objects.

There are different types of moments that may be used to solve the problem of image coding. These moments are different in their properties and their computation complexities.

Conventional Moments

The conventional definition of the two-dimensional moment of order (p + q) of a function f(x, y) is given by:

$$M_{pq} = \Sigma\Sigma x^p y^q f(x,y)^r$$

$$p,q = 0,1,2,...$$

(6.1)

Where the summation is over the area of the image. If the input image is a binary image, then r is equal to 0, otherwise r is equal to 1. $f(x,y)$ represents the gray-level value of the pixel at point (x, y).

The conventional moment sequence (M_{pq}) is uniquely determined by $f(x, y)$ and (x, y) is uniquely determined by the sequence (M_{pq}). It is always possible to find all orders of moments for an image, but a truncated set may offer a more convenient and economical representation of the essential shape characteristics of an object.

The conventional moments of an object are changed by the primary characteristics of the object in the following manner:

   1. Scale change:

The conventional moments of $f(x, y)$ after a scale change by a factor λ are defined by:

$$M'_{pq} = \Sigma\Sigma x^p y^q f(x/\lambda, y/\lambda)$$

(6.2)

   2. Translation:

The moments of $f(x, y)$ translated by an amount (a, b) are defined by:

$$M'_{pq} = \Sigma\Sigma (x+a)^p.(y+b)^q f(x,y)$$

(6.3)

   3. Rotation:

The moments of $f(x, y)$ after a rotation by an angle θ about the origin are defined by

$$M'_{pq} = \Sigma\Sigma (x\cos(\theta) + y\sin(\theta))^p.(y\cos(\theta) - x\sin(\theta))^q f(x,y)$$

(6.4)

   4. Reflection:

A reflection about the y-axis results in a change of sign of all moments, which depend upon an odd power of y. i.e.,

$$M'_{pq} = \sum\sum x^p y^q f(x, -y).$$ (6.5)

For object recognition and image analysis, a truncated set of moments is needed. A complete moment set of order n is always defined to be a set of all moments of order n and lower. A complete moment set of order n contains (n + 1)(n + 2)/2 elements.

Rotational Moments

The rotational moments for an image f(x, y) are defined by:

$$FnL = \int_{-\pi}^{\pi} \int_{0}^{\infty} r^n \exp(iL\theta) f(x,y) r dr d\theta$$ (6.6)

Where

$$r = \sqrt{x^2 + y^2}$$

$$\theta = \tan^{-1}(y/x)$$

$n =$ order of the moment.

$$n \geq L \geq 0$$

A truncated set of rotational moments of order n contains the same information as a truncated set of conventional moments of the same order.

The rotation normalization of these moments is very simple, but the translation and scale normalization is very difficult and time consuming. The standard rotational moments can be obtained by normalizing the rotational moments with respect to scale, translation and rotation.

Orthogonal Moments

There are two types of orthogonal moments:

   a. Legendre moments:

These moments are based on the conventional moments. The legendre moment of order (p + q) is defined by:

$$L_{pq} = \frac{(2p+1)(2q+1)}{4} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} Pp(x) Pq(y). f(x,y). dxdy$$ (6.7)

Where Pp(x) and Pq(y) are Legendre polynomials. Legendre polynomials are defined by:

$$Pn(x) = \sum_{k=0}^{n/2} (-1)^k \frac{(2n-2k)!}{2^n k!(n-k)!(n-2k)!} x^{n-2k}$$ (6.8)

Advantages:

   1. The inverse transform to obtain f(x, y) from the sequence (L_{pq}) is very simple because of orthogonality property.

2. They are more suitable for optical methods than the conventional moments, since the range needed to record Legendre moments is less than that needed for conventional moments.

Disadvantages:

Much computation is required to normalize the Legendre moments with respect to scale, translation and rotation.

b. Zernike Moments:

These moments are defined based on rotational moments. The Zernike moment of order n is defined by:

$$A_{nL} = \frac{n+1}{\pi} \int_{-\pi}^{\pi} \int_{0}^{\infty} \left[V_{nL}(r,\theta) \right]^* f(x,y).rdrd\theta \tag{6.9}$$

Where $[V_{nL}]^*$ is the complex conjugate of Zernike polynomial. The Zernike polynomials are defined by:

$$V_{nL}(r,\theta) = \sum_{s=0}^{(n-1)/2} (-1)^s \frac{(n-s)!}{s! \left(\frac{n+L}{2} - s \right)! \left(\frac{n-L}{2} - s \right)!} r^{n-2s} \exp(iL\theta) \tag{6.10}$$

And r is same as before.

The rotation normalization of Zernike moments is very simple but the translation normalization is very difficult. The inverse transform of f(x, y) from a sequence $\{A_{nL}\}$ is very simple because of orthogonality property.

Object Classification using Moments

A complete moment set can be created by any type of the moments discussed above. However, the conventional moments require minimum computation time for scale and translation normalization. The only operation to be considered is the rotation normalization. Simple parallel hardware architecture may offer a significant speedup to the rotation normalization process. Hence, moment invariant and standard moments are considered for object classification.

Fourier Descriptors

The Fourier descriptor method is a well-known method of describing the shape of a closed figure. Depending on the technique used to calculate the FD's, the FD's can be normalized to be independent of the translation, the rotation and the scale of the figure.

The two main steps involved in getting the FD's of a digital picture are:
1. The transformation of a binary matrix to a polygonal boundary.
2. The calculation of the FD's of the closed curve.

Calculation of Fourier Descriptors

There are different techniques used to calculate the FD's of a closed curve. The two most common techniques are discussed below.

a. The technique suggested by Cosgriff and Charles T. Zahn:

This technique involves many steps to calculate the FD's as discussed below.

1. A starting point on the boundary is selected.
2. The boundary is traced in the clockwise direction and at each pixel the last and next neighbors are found and noted down.
3. The vertices of the curve are found.
4. The length (ΔL) between any two successive vertices is calculated. The length between the starting point, and a vertex V_i is denoted by L_i, and it is equal to $\sum_{k=1}^{i} \Delta L_k$
5. The perimeter is calculated by using the following formula:

$$L = \sum_{i=1}^{m} \Delta L_i \qquad \text{where m is the number of vertices.}$$

6. The angular change ($\Delta \phi$) at each vertex is calculated.
7. Next, the Fourier series coefficients as shown below, are calculated.

$$a_o = -\pi - \frac{1}{L} \sum_{k=1}^{m} L_k \Delta \Phi_k \qquad (6.11)$$

$$a_n = -\frac{1}{n\pi} \sum_{k=1}^{m} \Delta \Phi_k \sin \frac{2\pi n L_k}{L} \qquad (6.12)$$

$$b_n = \frac{1}{n\pi} \sum_{k=1}^{m} \Delta \Phi_k \cos \frac{2\pi n L_k}{L} \qquad (6.13)$$

Where $\Delta \Phi_k$: is the angular change at vertex k.

L_k : is the arc length to vertex k.

L   : is the perimeter.

m   : is the number of vertices.

8. Lastly, the polar coordinates are calculated as follows:

$$A_n = \sqrt{a_n^2 + b_n^2}$$

$$\alpha_n = \tanh \frac{b_n}{a_n} \qquad (6.14)$$

The A_n's and the α_n's are the Fourier descriptors. It is usually sufficient to represent a closed curve by the first 10 harmonics. The calculated FD's are put in a vector. This vector is called the feature vector.

The FD's calculated by this technique are independent of rotation, translation and scale.

b. The technique used by Wallace and Wintz , and Kuhl et al.:

This technique retains almost all shape information, and is computationally efficient because it uses the FFT. The steps involved in this technique are explained below.

1. A starting point on the boundary is selected.
2. The original contour is sampled and then replaced by a piece-wise linear contour.

3. The piece-wise linear contour is represented by a chain code representation.
4. The length of the contour is computed.
5. The contour is uniformly resampled at spacing chosen to make the total number of samples a power of 2.
6. The FD's are computed by simply taking the FFT of the resultant samples.
7. The FD's should be normalized such that the contour has a standard size, rotation, translation, and starting point.

First, we set $A(O)$ equal to zero to normalize position. Second, scale normalization is accomplished by dividing each coefficient by the magnitude of $A(1)$. Third, rotation and starting point normalization are accomplished by making the phases of the two coefficients of largest magnitude to be zero.

The normalized FD's (NFD's) are put in a vector. This vector is called the feature vector and it is used to recognize the object.

To recognize an object, there should be a training set. The training set should contain all the possible feature vectors for the all-possible types of objects. The recognition of an unknown feature vector can be achieved by using one of the following two techniques:

    a. Distance-Weighted K-Nearest Neighbor Rule.
    b. Chi-Square Technique.

Because FD's are just independent of one orientation angle (θ_y), the training set should contain feature vectors for any given angles, θ_x and θ_z

Expert Systems

Expert systems have been successfully used in different fields of Engineering, as a substitution of the traditional methods.

An expert system is a system designed to accomplish a certain goal, by following the appropriate set of rules that is a part of a larger, more general group of rules, given by experts in the specific field.

Expert systems consist of two units, the knowledge base management unit and the knowledge base unit. The knowledge base management interfaces between the knowledge base and the outside world and hence facilitates to enter the basic knowledge and to construct this knowledge for efficient retrieval and decision-making.

Object representation

The knowledge base has two segments. One segment contains the specific facts, which define the problem, and the other segment contains the rules that operate on the facts during the problem solving process.

The knowledge in the knowledge base should be structured in a very powerful way to facilitate the searching process and to reduce the searching time. Three knowledge structures are common. These are:
1. Relational knowledge base structure, which takes a form of a table.

2. Hierarchical knowledge base structure or tree structure. Each parent bas one or more descendent, but each descendent has only one parent.

3. Network structure. Each parent bas one or more descendent and each descendent can have one or more parent.

It is easy to update the knowledge in a relational structure but searching takes a long time. The hierarchical structure reduces the searching time. The network structure allows many-to-many relation.

In general each knowledge base bas the following features:

1. A fact base which is equal to the conventional database.

2. A rule base that consists of production rules that operates on the fact base.

3. Confidence factors, which indicate the degree of confidence that, can be placed in facts and rules.

4. Meta-rules, which determine when and how the rules in the rule base are to be executed.

The knowledge base management unit accepts the extracted features of an object and then represents them in a way that is appropriate for symbolic processing. There are several ways to represent these features: first order predicate calculus, frames and semantic networks.

In the problem of recognizing a tank, for example, first-order predicate represents the fact that a gun is located over a small rectangle in the form [6]:

OVER (GUN, SMALL-RECTANGLE)

A frame for the same fact might be in this form

name of frame: F1
type of frame : OVER
object 1   : GUN
object 2   : SMALL-RECTANGLE

In the semantic network, each node contains an object and the lines between the nodes represent the relationship.

The knowledge base is unique to this particular problem. This knowledge base contains rules that show the relationship among several pieces of information. The most common rule is the production rule, such as:

IF THERE IS A GUN, AMD THE GUN IS OVER A RECTANGLE, AND THE RECTANGLE IS OVER WHEELS, THEN THERE IS A TANK.

Here we have three antecedents connected by logic AND, which when satisfied lead to the consequence that there is a tank [6]. When all the three antecedents are present as assertions, they match the left-hand side of the production rule, leading to the consequence there is a tank.

In addition, there are two further steps to the making of matches:

1. The production rules may be listed in some order to facilitate the matching process.

2. The search strategy unit may be able to invoke a sequence of rules that finds the match more quickly than random search lists.

The search strategy unit searches the knowledge base for problem solution or object identification. Knowledge base search is the foundation of artificial intelligence systems. The searching process depends on the knowledge structure or the arrangement of the production rules in the knowledge base. The simplest arrangement of the production rules is to list them sequentially with no order. Each asserted fact is then run through the production rules until the match is found. With a small number of rules, it is practical to search a random list of rules. In

3D-object recognition problem, there are many rules and hence the knowledge base is huge. For this type of problems, the knowledge structure must be designed properly to decrease the searching time.

<div align="right">Searching algorithms</div>

As stated before, the searching algorithm depends on the knowledge structure. Taking the best knowledge structure, the following searching algorithms may be applied:

A. Depth-first search:

This technique takes one branch of the knowledge structure tree and searches all the nodes of that branch to its last node. If a solution is not found, the procedure backs up to the next unsearched branch and continues the search to the last node and so on until a solution is found.

B. Breadth-first search:

This technique searches according to the level of a node. Level one will be the top node. The successors of this node will be labeled level two. This technique starts by searching all of the nodes on level two then level three and so on, until a solution is found.

C. Heuristic Search:

The above two searching techniques require the searching through the entire knowledge base in order to find the solution. When the knowledge base is huge, the time for solution search can become a real problem. The Heuristic Search searching technique is devised to reduce the searching time.

Heuristics is defined as any information that would bring closer to a problem solution. Such information often takes the form of rules-of-thumb that may be applied in decision-making processes.

Heuristic information can be applied at many nodes along the solution search path, and decision can then be made as to which branch of the knowledge structure is more likely to hold the last solution.

Heuristics are generally applied in one of the following two ways:

1. When deciding whether a particular branch of the knowledge structure should be eliminated.
2. When deciding which branch of the knowledge structure to be searched next.

The inference unit is used to obtain the object identity. The inference or reasoning process accepts the assertions from the input image, and searches the knowledge base for identification. Two approaches to inference are common: deduction inference and induction inference. Deduction or backward inference begins with an hypothesis or a goal, then proceeds to examine the production rules, which conform to the hypothesis. Induction or forward inference examines the assertions then draws a conclusion, which the conditions conform to. In other words, deduction starts with the goals, while induction starts with the facts or the assertions. Different 3D object recognition systems that utilize AI techniques has been built and tested. The performance of those systems is very optimistic.

Neural Networks

A neural network consists of a number of nonlinear computational elements called neurons or nodes. The neurons are interconnected with adaptive elements known as weights, and operate in parallel environment. The structure of neural networks is similar to simplified biological systems. Recently, neural networks have become a hot area of research, because they possess the following characteristics:

- Highly parallel structure, hence a capability for fast computing.
- Ability to learn and adapt to changing system parameters (e.g. inputs).
- High degree of tolerance to damage in the connections.
- Ability to learn through parallel and distributed processing.
- Nonparametric values, not dependent on as many assumptions regarding underlying distributions as traditional networks (e.g. classifier or optimizing circuits).

Neural networks have been used as associative content addressable memories (ACAM) and as classifiers. In an ACAM, the memory is associated not by an address, but rather by partially specified or many versions of the stored memory pattern.

A classifier assigns an input sample to one of the predetermined classes. The use of neural networks has recently been proposed for creating dynamic associative memory (DAM), which utilizes supervised learning algorithms, by recording/learning to store the information as a stable memory state.

The neural network minimizes the energy function formed by the mean square error constructed by the difference between the actual training signal and the signal estimated by the network.

Back propagation

Back propagation is one of the most popular supervised training methods for neural networks. It is based on the gradient-descent technique for minimizing the square of error between the desired output and the actual output. Following is a summary of the back propagation training algorithm for a two-layer neural network with n inputs, l outputs and j neurons in the hidden layer.

1. Store the training set $\{(X^k, d^k): k = 1, 2, 3, ..., n$. Where X^k is the input vector and d^k is the desired output vector.
2. Initiate V and W to small random numbers (between -1 and 1). Where V is the weight matrix for the hidden layer ($j \times n$) and W is the weight matrix for the output layer ($l \times j$).
3. For each pair (X, d), compute the foreward pass as follows:

- Compute net_j, $j = 1, 2, ..., n$. Where $net_j = X^T V$ (7.1)

- Compute Zj, where, for a unipolar sigmoid activation function, Zj can be given by:

$$Zj = \frac{1}{1 + e^{-\lambda net_j}} \tag{7.2}$$

- Compute net_l, $l = 1, 2, ..., L$. Where $net_l = z^T W$ (7.3)

- compute y_l where for the same unipolar activation function, y_l is given by:

$$y_l = \frac{1}{1 + e^{-\lambda net_l}} \qquad (7.4)$$

4. compute the backword pass as follows:

- $\delta ol = (d_l - y_l) f_o'(net_l),\ l = 1, 2, ..., L.$
- Compute $\Delta W_{lj} = \eta \delta ol Zj$. Where η is the iteration step. $\qquad (7.5)$
- $\Delta V_{ji} = \eta f'(net_j) X_j \sum\limits_{l=1}^{L} \delta ol W_{lj}$ $\qquad (7.6)$

5. Update weights as follows:

- $W_{lj}^{k+1} = W_{lj}^{k} + \Delta W_{lj}$ $\qquad (7.7)$
- $V_{ji}^{k+1} = V_{ji}^{k} + \Delta V_{ji}$ $\qquad (7.8)$

6. Check convregance by comparing the resultant output with the desired ones. If not converged, goto step 3.

Figure (16) shows a schematic for a two-layer neural network.

Neural Networks for Image Compression

Recently, neural networks have been used for image compression. Artificial neural network models called connectionist models or parallel distributed processing models have received much attention in many fields where high computation ratios are required. Many neural network approaches for image compression yield performance superior to that of the discrete traditional approaches. The estimated output is compared with the actual one for learning. For back propagation algorithms, image data compression is presented as an encoder problem. In fact, the weight matrices encode the principal components of the given image, i.e., after convergence it decomposes the given image into an eigenvalue of decreasing variance. Back propagation is so powerful for image compression that singular value decomposition techniques will be quite appropriate. Figure (17) shows one possible application for neural network in image transmission. Digital image generated with an 8-bit imagery system may be reduced in size by feeding it to an 8-to-4 multi-layer neural network. On the other end of the line, the compressed image is processed by a 4-to-8 neural network for reconstruction. Compressing images before transmitting them speeds up the transmission time substantially.

Neural network is an ideal tool for pattern recognition. Any recognition system needs to be trained to recognize different patterns, and training is the most important part in the design of neural network systems. All pattern recognition systems try to imitate the recognition mechanism carried out by humans. Since neural network is a simplification of human neural system, it is more likely to adapt to the human way of solving the recognition problem than other techniques and systems. Finally, we can look at pattern recognition as a classification problem, which is best handled by neural network systems.

The design of neural network system for pattern recognition starts with collecting data on each of the objects that is to be recognized by the system. A class is assigned to each object, and the collection of data and classes is used to train the system.

FILENAME.APP=4106FG01.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (1) Cosine function with amplitude A and frequency 1 Hz.
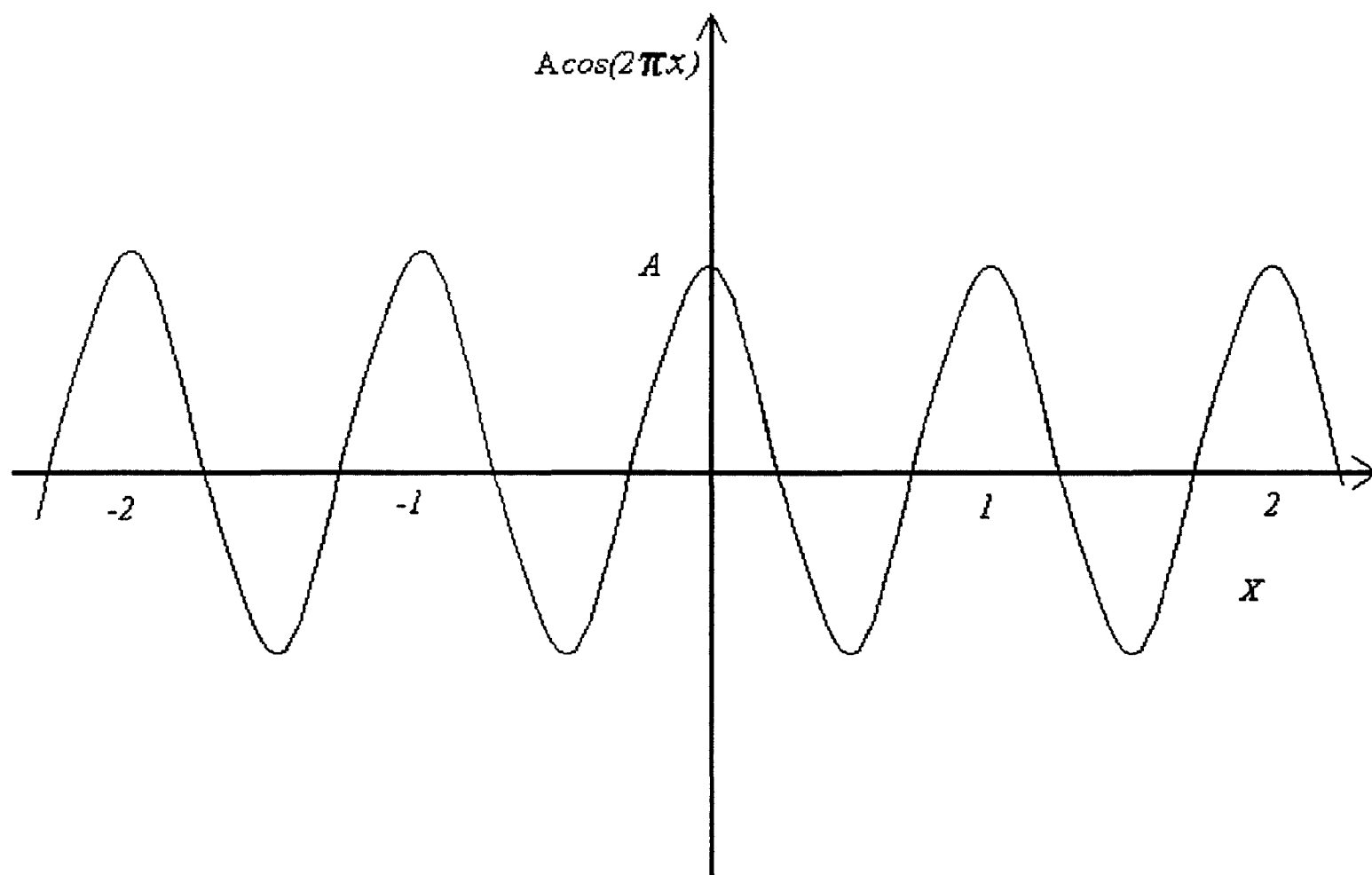
Figure (2) Power spectrum of the cosine function with amplitude A and frequency 1 Hz.
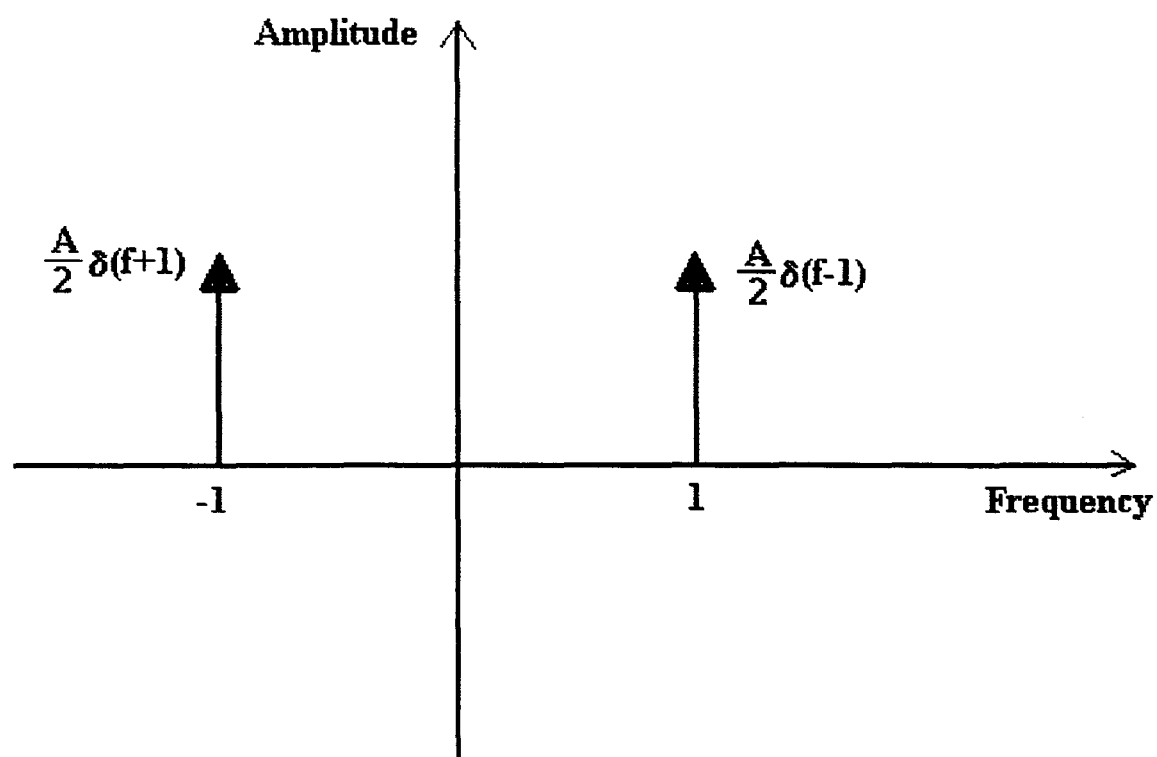
FILENAME.APP=4106FG03.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (3) Truncated Cosine function.
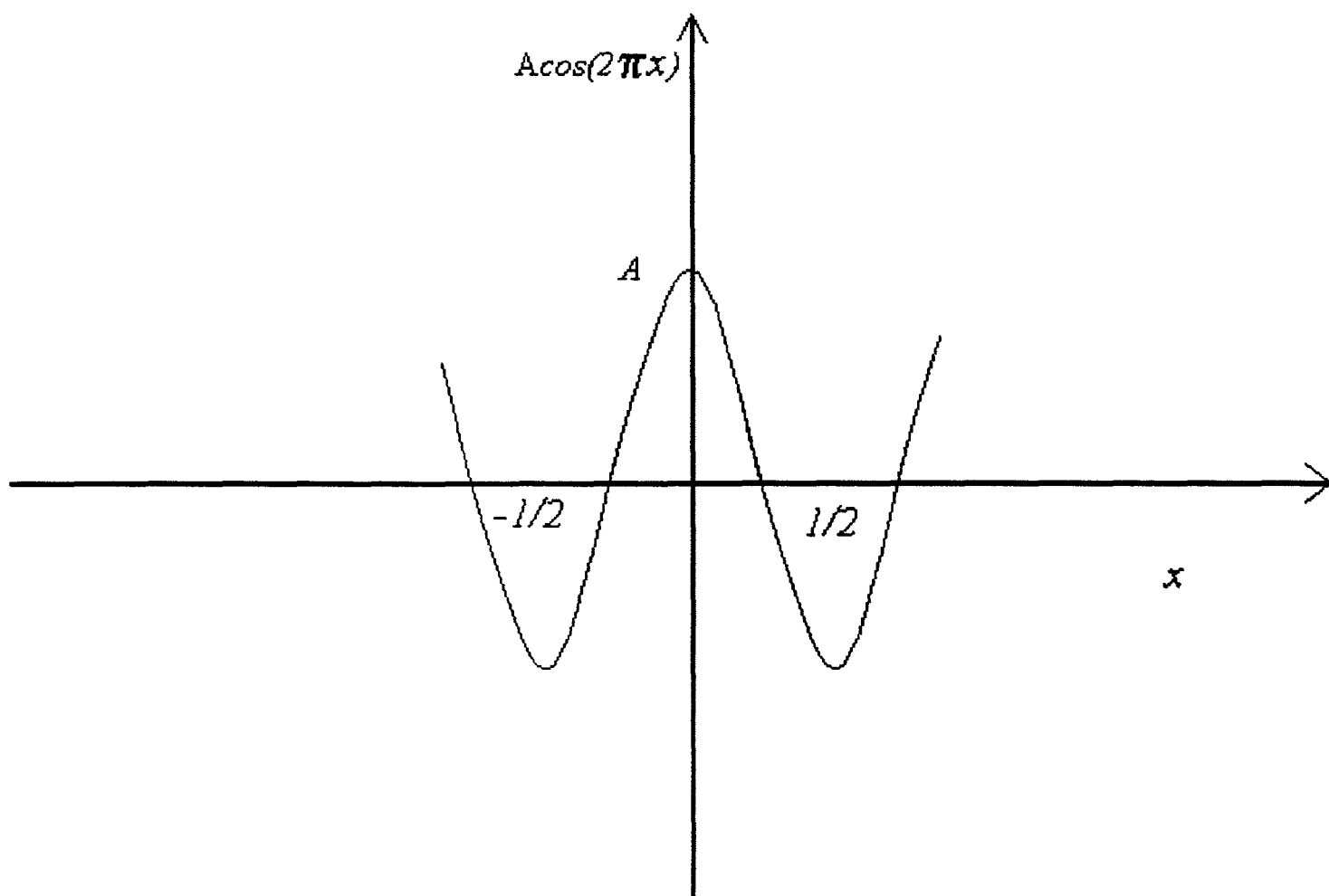
FILENAME.APP=4106FG04.PSD
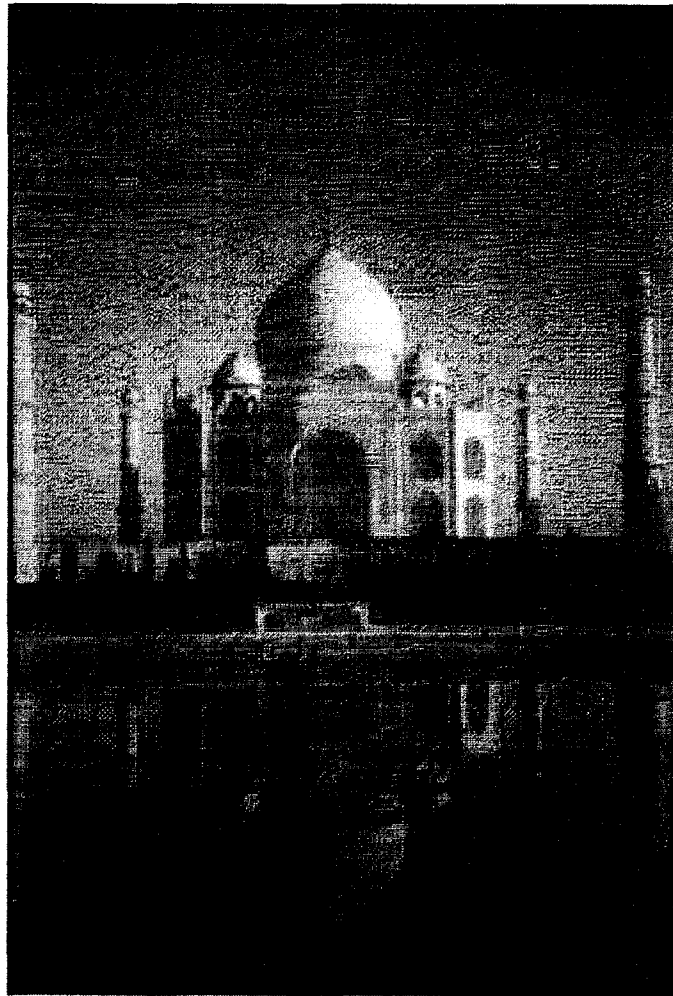APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (4) Power spectrum of a truncated cosine function with amplitude A and frequency of 1 Hz.
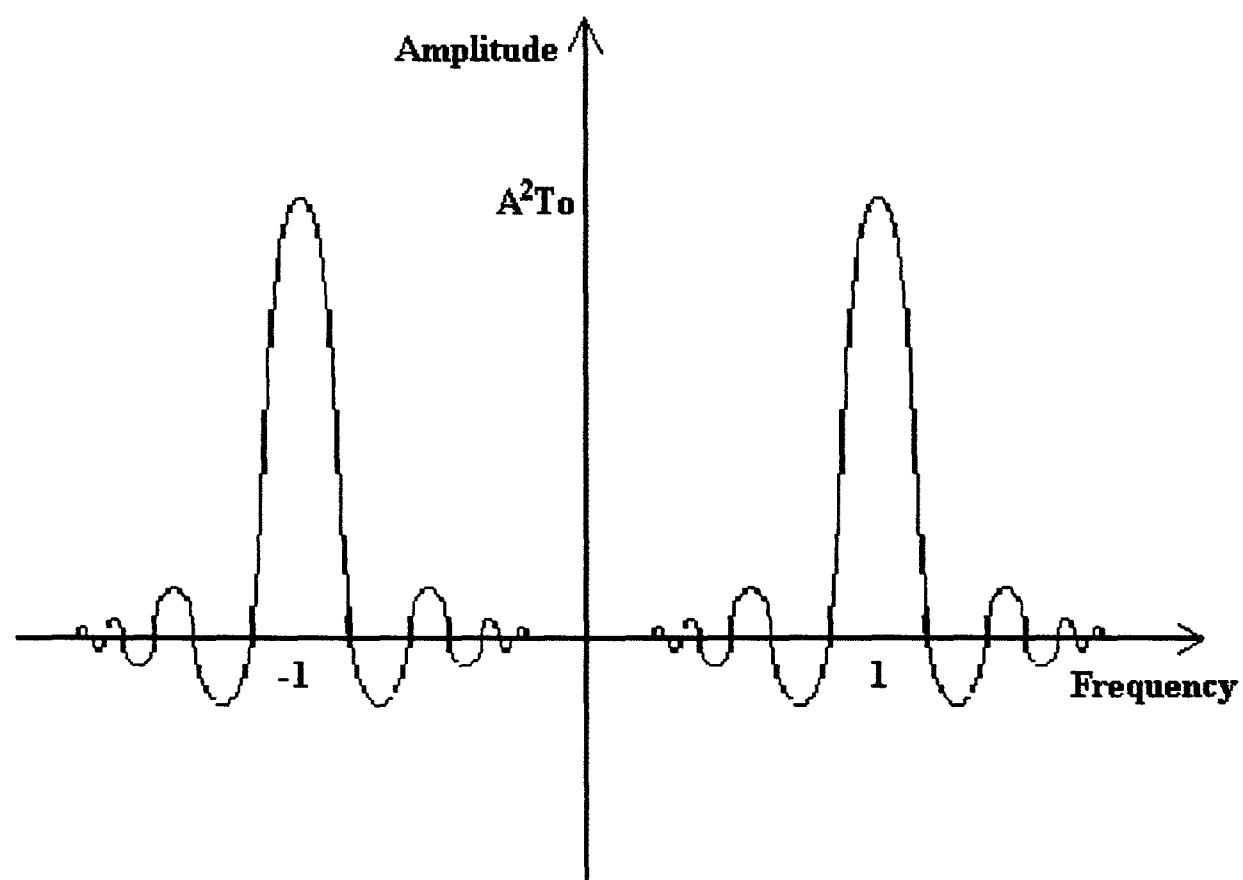
Figure (5) Antialiasing filtering (a) Square signal with higher frequency noise.
(b) Digitized signal with noize.
(c) Continuous signal after using 2nd. Order Bessel filter.
(d) Digitized filtered signal.
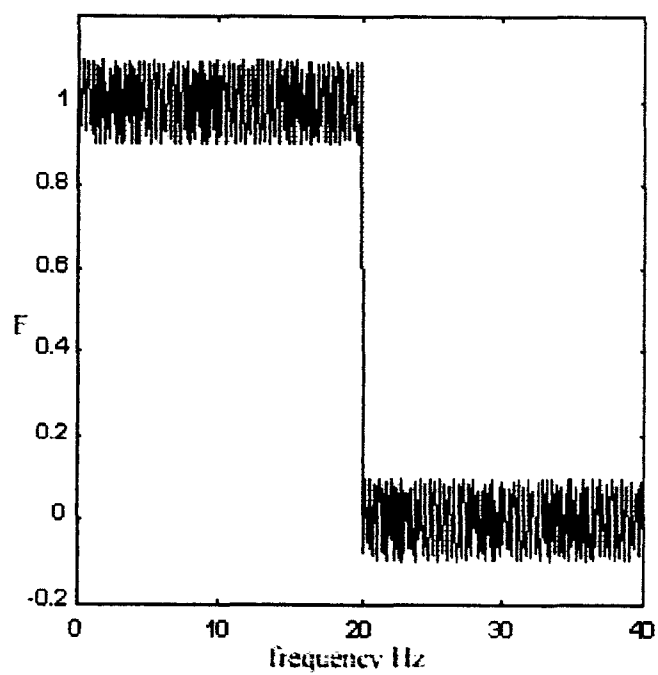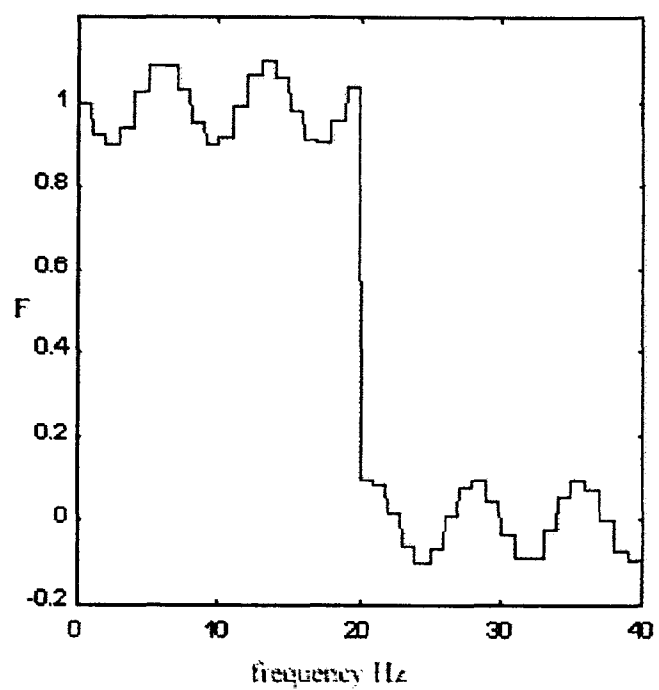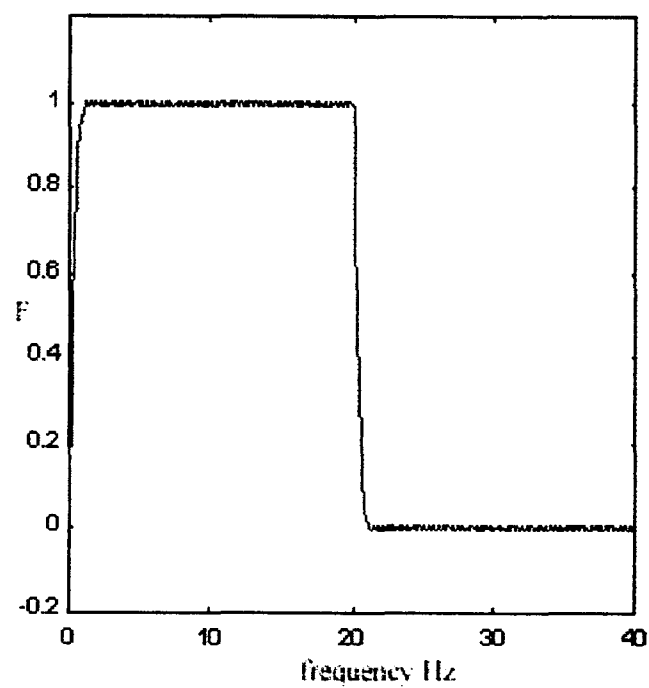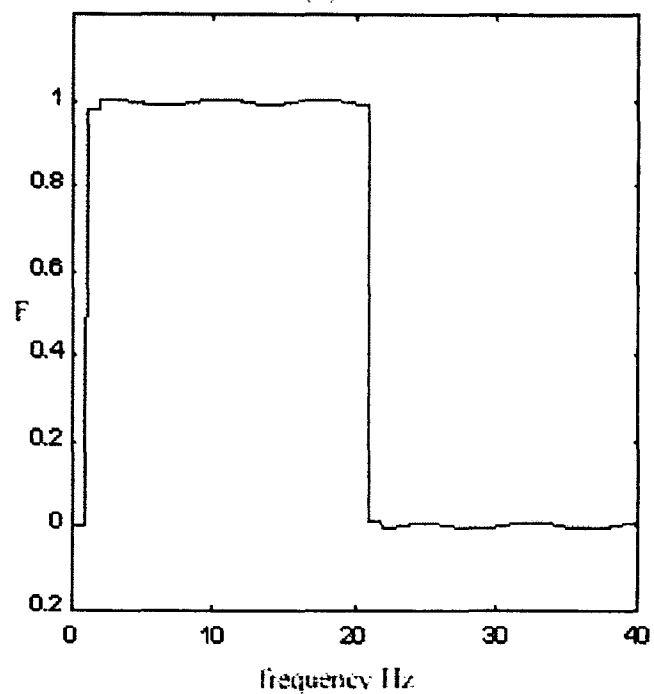
Figure (6) Simplified model for degradation process.
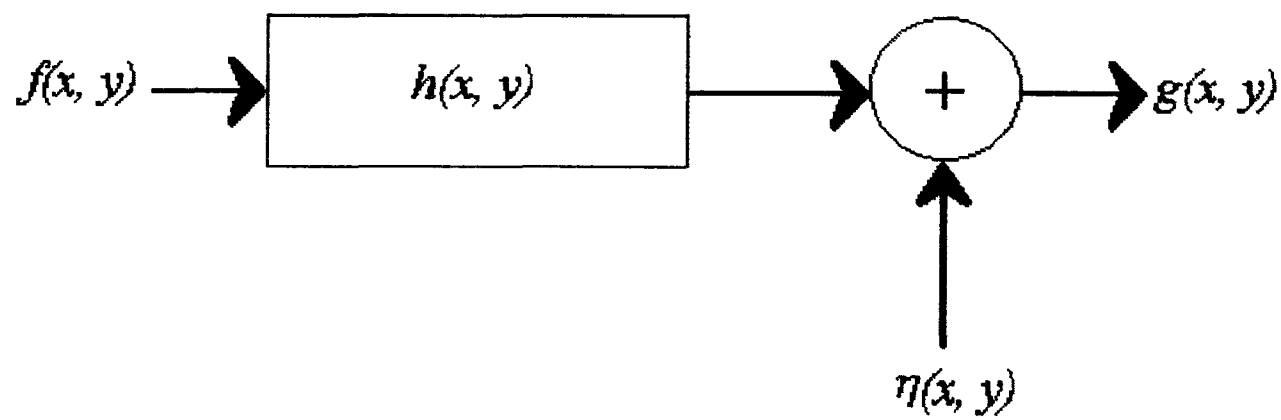
Figure (7) Typical spectra of deconvolution filtering image restoration system.

cut-off frequency

Theoretical response

frequency (Hz)

Actual response

frequency (Hz)

Inverse response

frequency (Hz)

Restored response

frequency (Hz)
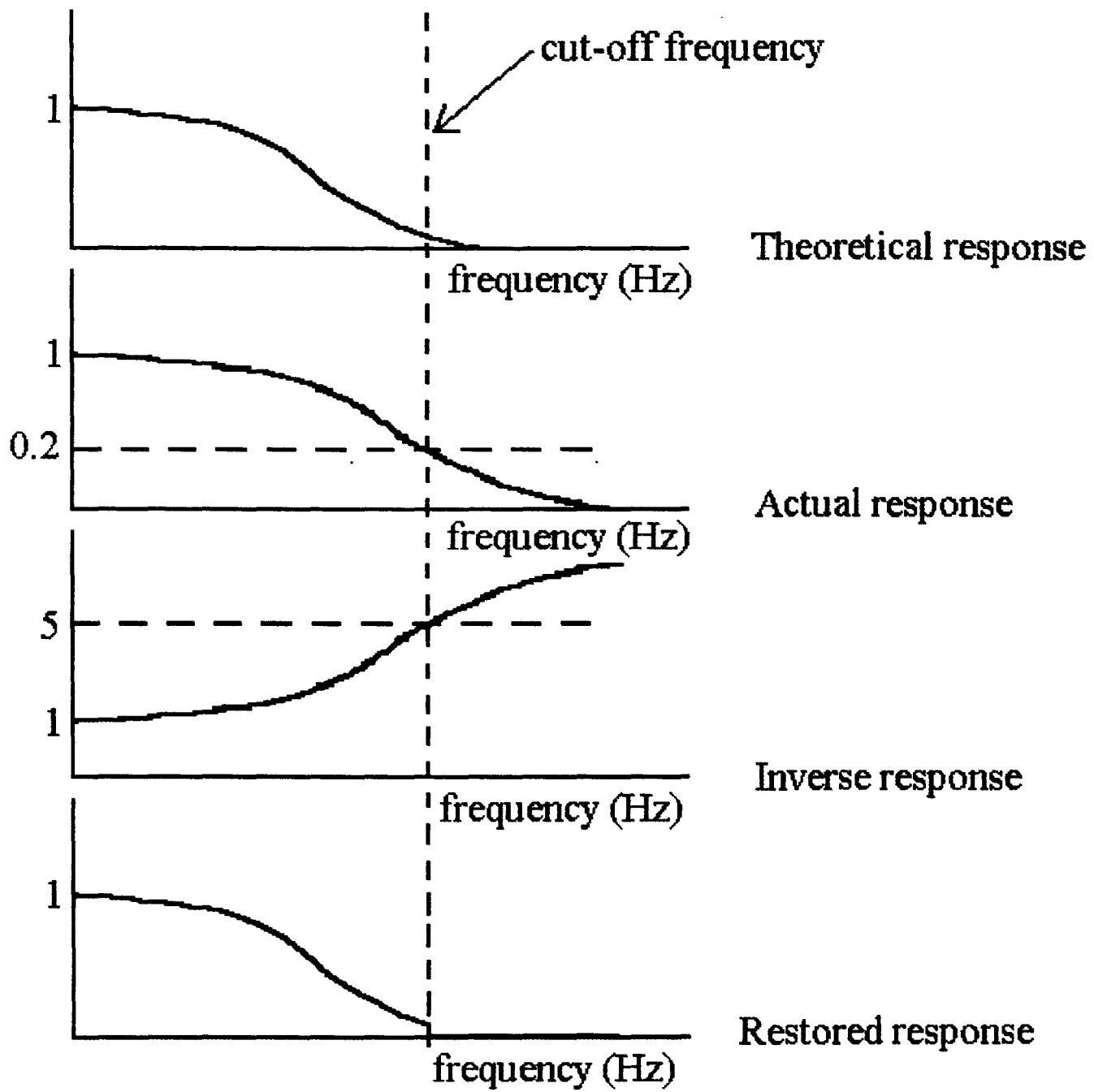
FILENAME.APP=4106FG08.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (8) Linear point operator.

N(Gi)

255

$N(G_i) = aG_i + b$

a

G_i 255
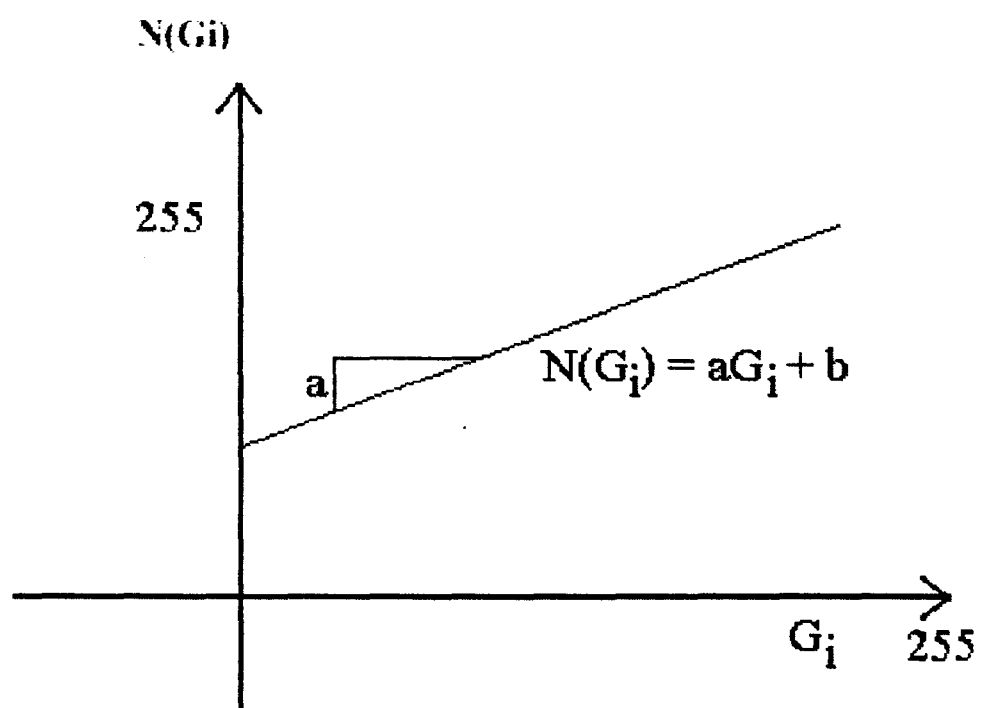
Figure (9) Effect of window and level technique on the image histogram.

Output Histogram

L

W

Original Histogram
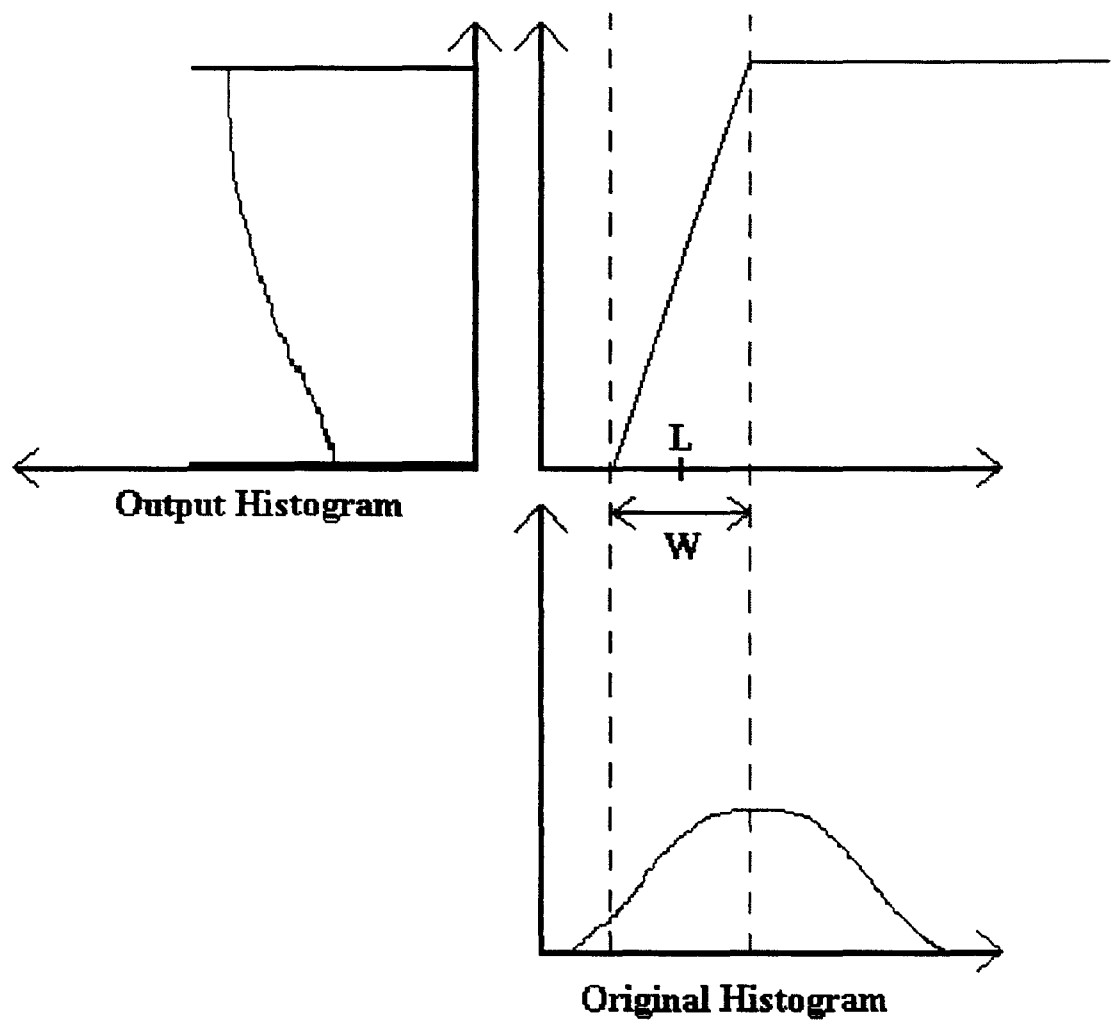
FILENAME.APP=4106FG10.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (10) Photo of Taj-Mahal with arbitrarily noise.

Figure (11) Low-pass filtering of the image removes most of the noise but blurs the image.
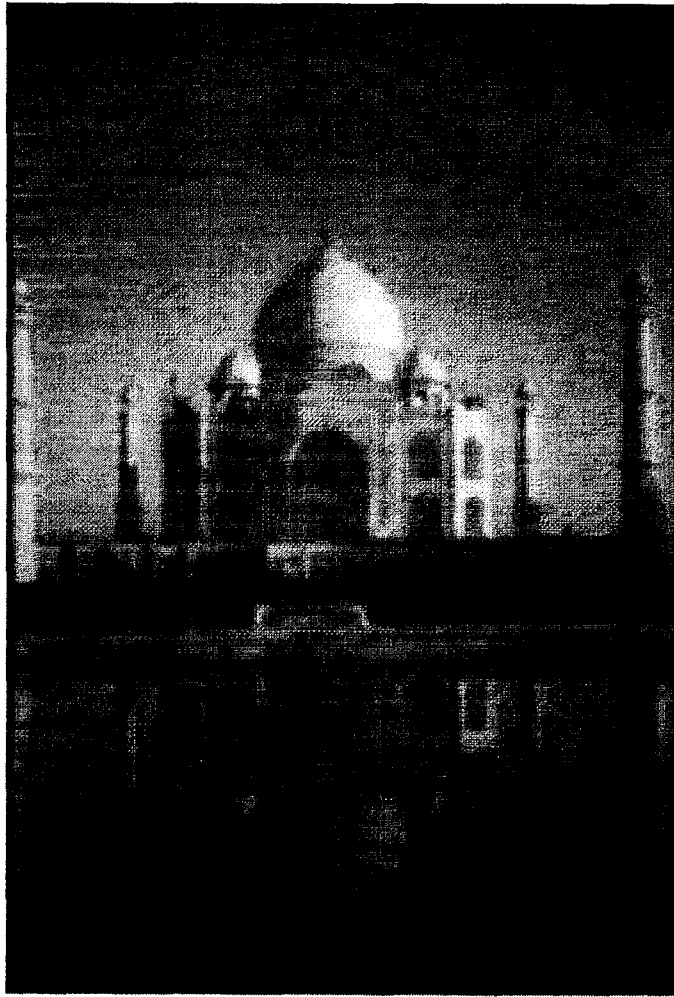
Figure (12) High-pass filtering of the image sharpens the edge details but it increase background noise signal power.
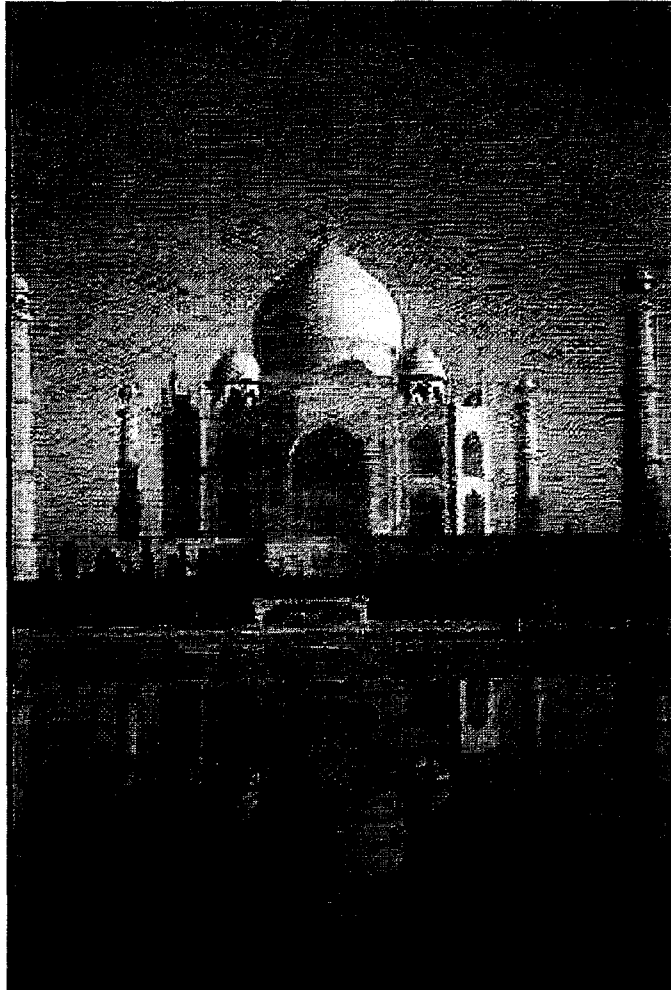
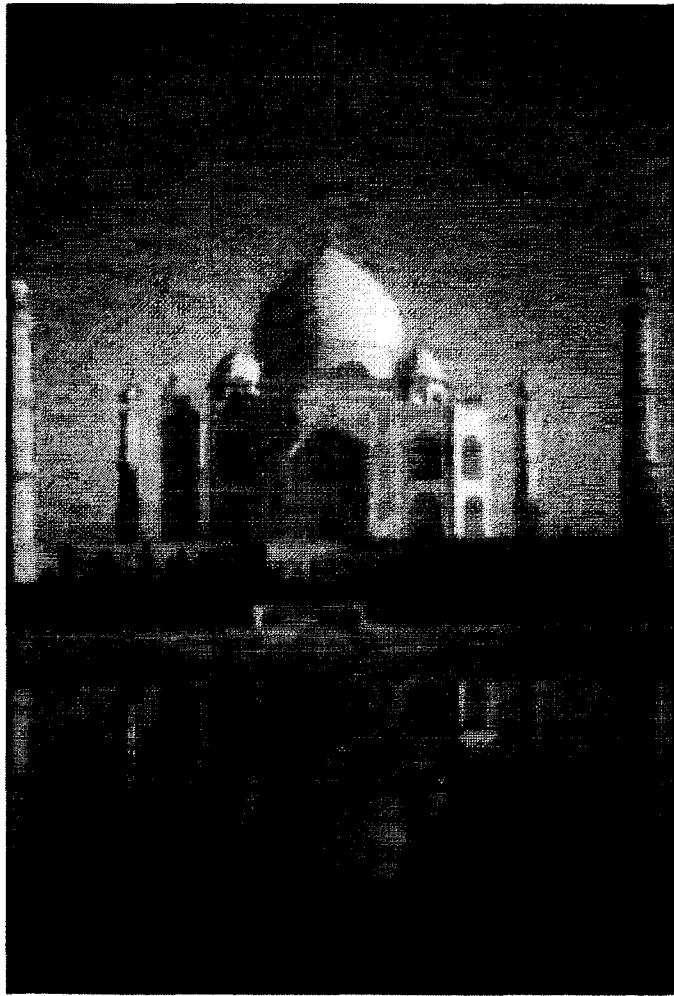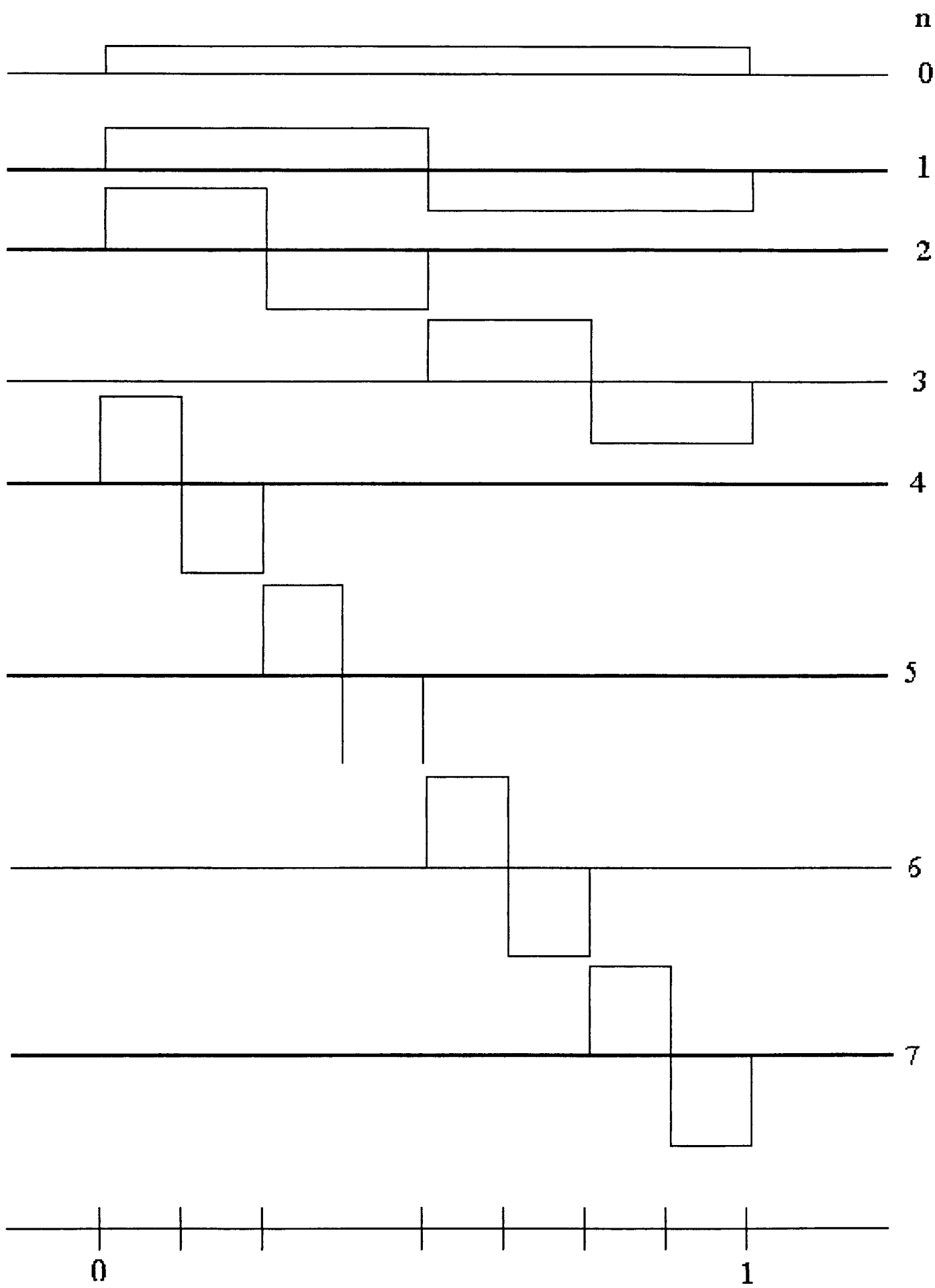Figure (13) Median filtering reduces random noise while enhancing edges.

FILENAME.APP=4106FG14.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (14) The use of wavelets to process raw data.

Raw Data → Wavelet decompisition → Threshold → Wavelet Compression → Processed Data

FILENAME.APP=4106FG15.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (15) The Harr transform basis functions.

Figure (16) Two-layer neural network with n inputs, j hidden neurons and 1 outputs.

v

w

1

$z1$

2 $z2$

3 $z3$

4 $z4$

j zj

$x1$

$x2$

$x3$

x_n

1

1

1 →$y1$

2 →$y2$

3 →$y3$

l →y_l

Inputs

Hidden Layer

Output Layer

Outputs
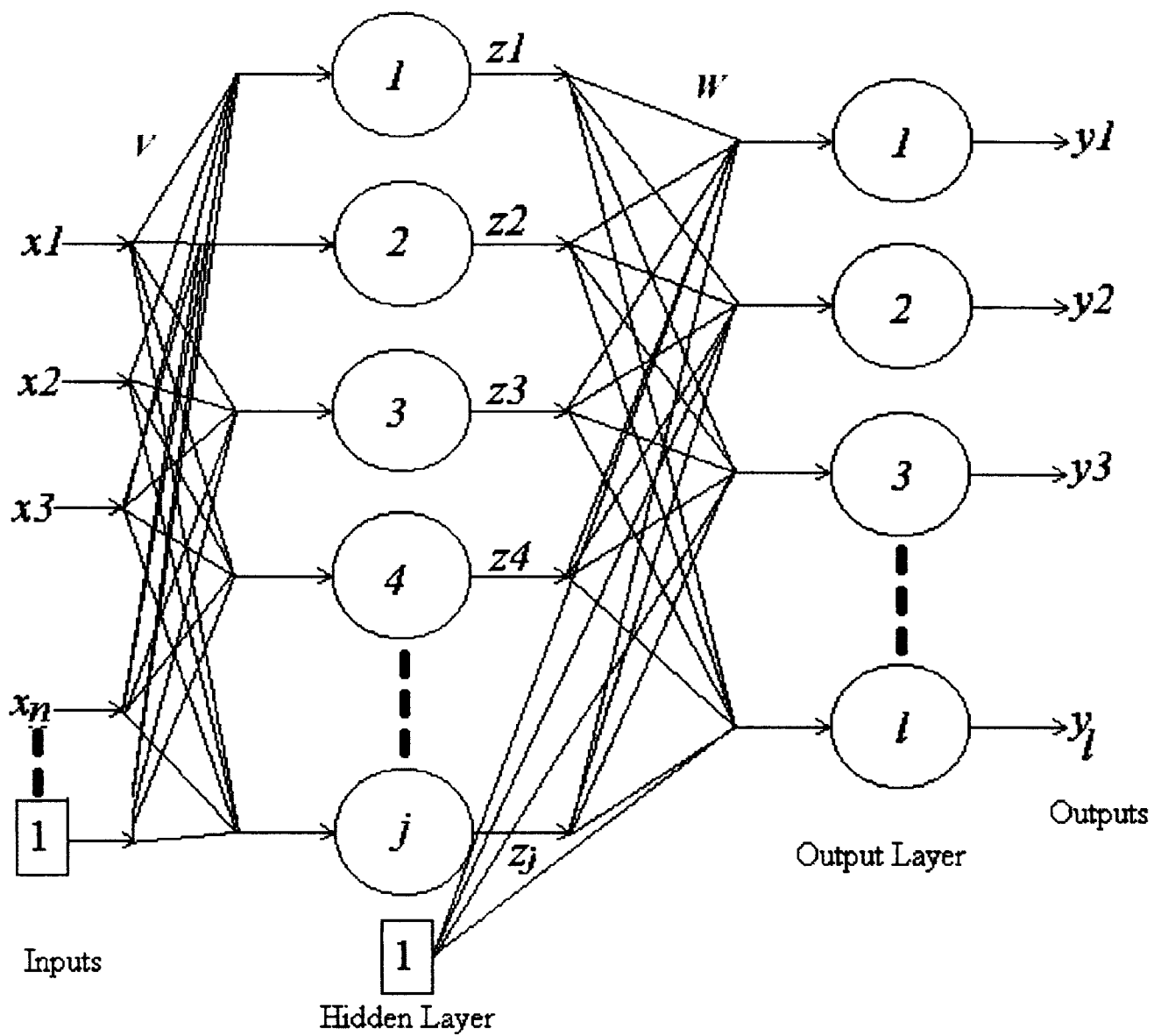
FILENAME.APP=4106FG17.PSD
APPLICATION: ADOBE PHOTOSHOP VERSION 4.0

Figure (17) The use of neural networks for image compression and transmission.

8 bit input

From
Image

Multilayer
neural
network with
8 inputs and
4 outputs

Transmitting
end

4 bit output

Multilayer
neural
network with
4 inputs and
8 outputs

Receiving
end

8 bit output

To display